Artur Sogomonyan

arthur.sogomonyan@gmail.com | artur-sogomonyan | GitHub
+43-677-641-145-81 | Austria | Portfolio Website

Education

Data Science — *Master of Science* 2022 – now
Technische Universität Wien, Austria

Information technologies and telecommunications — *Bachelor of Science* 2014 – 2020
Bauman Moscow State Technical University
Diploma thesis "3D printer failure detection system based on computer vision detection"

Work Experience

Data Scientist in Project Rail4Future — *Technische Universität Wien* Feb 2024 – Jul 2024
- Analyzing rail irregularities, refactoring and optimization of the LSTM models. Working with industrial partners Siemens and ÖBB.
- Performed significance tests, achieved robustness, and dramatically improved the performance of the model.

CTO, Data Scientist — *susteam* Jan 2023 – present
- Architecture and Full Stack development of the ESG system.
- Developed a pipeline with OCR, information retrieval, Data Analysis and Machine Learning. RAG implementation for LLMs.
  FastAPI, Amazon Web Services (AWS), Amazon EC2, S3, React.js, Bert, LLama3

Machine learning analyst — *GLOSAV* Feb 2021 – Jul 2021
- Engineered a sophisticated algorithm to detect railway routes with object detection and tracking mechanisms. Developed a robust way to track obstacles on the rails. Achieved good performance and latency for real-time stream less than 5ms.
- Contributed to the design and implementation of an expansive microservice architecture of the train autopilot system.
  Python, YOLOv5, OpenCV, NumPy

Project Manager, 3D-Modeller, Software Engineer — *Arterra Miniature* Feb 2020 – Mar 2021
- Working with customers, 3D visualization, prototyping, 3D printing, programming microprocessors and microcontrollers, work from scratch to delivery of projects.
  Blender, Cura, Arduino

Frontend Developer — *KRK Group* Sep 2019 – Dec 2019
- Developed corporate educational platform, React application, implementation of user interface, communication with databases.
  JavaScript, React

Full-Stack Developer — *Linwood* Aug 2018 – Sep 2019
- Implemented all stages of the online store development, subsequent testing, promotion and expansion of functionality for the required business tasks.

Web developer — *Freelance* Mar 2017 – Aug 2018
- Carried out projects, as Front-end and Back-end. Implemented from scratch or refactored existing solutions for functionality extensions.
  PHP, Yii2, html, scss, Python, Django

Skills

Software Development
- docker
- Python
- C++
- R
- Django
- Flask
- FastAPI
- JavaScript
- React, Svelte

Machine Learning
- scikit-learn
- Pandas
- NumPy
- Tensorflow
- PyTorch
- OpenCV
- YOLO
- NLTK, SpaCy
- Bert, LLama3, GPT

Databases and Cloud
- Git
- PostgreSQL
- MongoDB
- Redis
- RabbitMQ
- nginx, Gunicorn
- Spark
- AWS EC2, S3
- Linux/Debian/Fedora

Languages
- English (C1)
- German (A2)
- Russian (C2)
- Armenian (B2)

Projects and Awards

AIM Hackathon - TIMETOACT Group — *2nd Place Winner* Oct 2024
- Developed a platform that uses Retrieval-Augmented Generation (RAG) to identify greenwashing in ESG reports. The system analyzed text from ESG reports, matched it with labeled data on greenwashing indicators, and provided a score based on the trustworthiness of the content. I also integrated an explainability feature to retrieve the most relevant quotes for identified greenwashing examples.

LangChain, OpenAI, FAISS, Python, Flask

LawSyncr — *Full-stack and Machine Learning Development* Aug 2024
- Conducted in-depth research and analysis to evaluate innovative ideas, leading to the creation of a comprehensive pitch deck.
- Led the end-to-end full-stack development of the project, including the implementation of a Minimum Viable Product (MVP) with integrated machine learning components.

Flask, Svelte, RAG, LLM

Arterra — *Portfolio Website Development* Aug 2024
- Designed and developed a portfolio website, integrating continuous integration/continuous deployment (CI/CD) pipelines for seamless updates via GitHub Pages.
- Utilized Obsidian vault data to create an interactive front-end that visualizes my professional journey using graphology and sigma libraries.

Svelte, graphology, sigma

TradeBot — *Crypto Trading Analysis and Future Trading* June 2024
- Engineered a bot to systematically scrape and collect trade data from Binance, allowing detailed real-time market analysis.
- Designed and implemented a robust trading strategy, optimizing both short-term gains and long-term trends.
- Automated the execution of trading strategies using a custom built bot, ensuring consistent and timely trades.

asyncio, FastAPI, BinanceAPI, pandas, mplfinance

DAVU Project TU Wien — *Document Analysis and Optical Character Recognition (OCR)* June 2024
- Conducted a comprehensive layout analysis of historical documents, identifying key structural elements to improve text extraction accuracy.
- Implemented an OCR pipeline using Tesseract, optimizing the process by experimenting with different binarization techniques, which resulted in significant improvements in Word Error Rate (WER) and Character Error Rate (CER).
- Analyzed the impact of various preprocessing methods, such as binarization and cropping, on OCR performance, achieving optimal text recognition by fine-tuning OCR engine parameters.
- Integrated advanced Named Entity Recognition (NER) using SpaCy and fine-tuned language models (e.g., LLaMA-3) to extract metadata from images, enhancing the accuracy of location, date, and photographer identification in historical datasets.

Tesseract, SpaCy, LLaMA-3, Python

Cryptex — *Full-stack Development of the NFT Platform* Feb 2024
- Implemented a secure authentication system using MetaMask, which allows users to store and manage seed phrases as encoded NFTs.
- Developed the platform's backend to interact seamlessly with the Polygon blockchain, ensuring efficient and secure transactions.

Svelte, Polygon, Solidity

Coding Austria — *Lip Reading app at Hackathon* Jul 2023
- Utilized Audio-Visual Hidden Unit BERT model for lips reading task for the Austrian Red Cross challenge.
- Integrated the model into the web app, passing the video stream from the smartphone to the model, reading the lips and generating a text passed as a message to the chat.
- This application allowed the Red Cross team to communicate in a very noisy environment, when it is not possible to use walkie-talkies.

AV-HuBERT, FastAPI, Next.js

Tourism Technology Hackathon — *2nd Place Winner* Nov 2023
- Implemented a data pipeline mapping data from different sources and building a knowledge graph for different tourism subjects. The result I visualized as an interactive map with the chat bot, that could wrangle the the point of interest and conclude the insights of the analysis.

Selenium, geoandas, folium

Studies and Publications

Publication — *IEEM24-F-0298* 2024
Comparative Analysis of Machine Learning-based Surrogate Modeling Approaches for Multi-body Dynamic Simulation in Railway Digital Twin Platform

Moscow Institute of Physics and Technology — *Deep Learning School* 2020 – 2021

Higher School of Economics University and Yandex — *School of Data Analysis* 2020 – 2020